

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-207487

(43)Date of publication of application : 28.07.2000

(51)Int.Cl.

G06K 9/03

G06F 3/00

G06K 9/00

(21)Application number : 11-011953

(71)Applicant : SONY CORP

(22)Date of filing : 20.01.1999

(72)Inventor : NAKAJIMA SHINJI

SHIONO TOMOKI

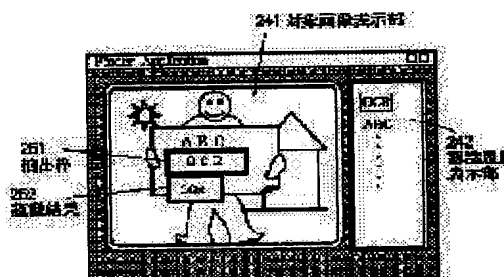
SUEYOSHI TAKAHIKO

(54) DEVICE AND METHOD FOR INFORMATION PROCESSING AND PROVIDING MEDIUM

(57)Abstract:

PROBLEM TO BE SOLVED: To simply and surely input a prescribed command to a personal computer.

SOLUTION: A CCD video camera picks up characters displayed on a prescribed object and overlaps and displays recognized results 252 obtained as a result of performing character recognition of the picked up image on an object image displaying part 241. An extraction frame 251 showing the range of image data utilized for the recognition is shown in the neighborhood of the results 252. Thus, it is possible for a user to intuitively grasp which part is extracted from the entire image being a recognition object to obtain the recognized results.



[0018] At the front of the display section 3, LCD (Liquid Crystal Display) 21 is disposed, and on its upper end portion, an image pickup section 22 is disposed in a freely rotating manner with respect to the display section 3. That is, this image pickup section 22 is configured to be able to rotate to an arbitrary position within the range of 180 degree between a direction identical to LCD 21 and a direction opposite to it (back face direction). A CCD video camera 23 is attached to the image pickup section 22.

[0045] At Step S4, the finder application 54C judges whether or not a user specified a character region. In the case where the character region is not specified, the process advances to Step S5, and the finder application 54C extracts a predetermined range in an image of one frame component stored in RAM 54, as a character region. Then, at Step S6, the finder application 54C transfers image data in the extracted character region to the character recognition engine 89 to have it recognized characters. The character recognition engine 89 has stored a wide variety of character patterns in advance, and compares the patterns with the image data within the range of the transferred character region, and recognizes whether or not a character is included therein. Then, the character

recognition engine 89 outputs the recognition result (character) to the finder application 54C.

[0046] In this manner, usually, the finder application 54C executes continuously the character recognition processing to an image of each frame, but it is also possible to specify a character region by a user operating the track point 5 etc. In the case where the user specified a character region, processing of Step S5 is skipped, and within the range of the specified character region, the character recognition engine 89 executes the character recognition processing.

[0047] At Step S7, the finder application 54C judges whether or not a character was recognized, on the basis of the output of the character recognition engine 89, and in the case where it is not recognized, the process returns to Step S2, and processing starting from the step is executed repeatedly.

[0048] At Step S7, in the case where it was judged that a character was recognized, the process advances to Step S8, and the finder application 54C displays a frame indicating the range of the character region in which a character was detected on GUI. That is, at this time, the finder application 54C controls the graphic chip 81 to display the character region in which a character was recognized, on the target image display section 241 of GUI, by means of the extraction frame 251. By this means, a user can know that a character was recognized from an image in which region. Furthermore, at Step S9, the

finder application 54C displays the character recognition result (character) 252 supplied from the character recognition engine 82 under the extraction frame 251. Meanwhile, at this time, in the case where plural candidates exist as the character recognition result, the candidates are displayed at the same time. A user selects a correct character from among the character recognition result 252 by operating the track point 5 etc., and in the case where one correct character has been already displayed, a user carries out an input operation for fixing the recognition result. In the vicinity of the character recognition result 252, the extraction frame 251 for indicating a range in which it was recognized is displayed, and therefore, a user can judge whether or not the recognition result is correct, promptly and surely.

[0052] The finder application 544C further judges, at Step S13, whether or not a predetermined one was selected from among the character recognition result displayed on the recognition history display section 242. That is, a user can select a predetermined one from among characters displayed on the recognition history display section 242 by operating the track point 5 etc. When a predetermined character was selected, the finder application 54C executes, at Step S14, processing

corresponding to the selected character.

[0053] For example, the finder application 54C judges, if the selected character string is a character string beginning by "http://", that it is URL (Uniform Resource Locator), and activates the Internet browser program 56E to have it accessed the URL. In the case where the selected character string is a character string separated by "@" and ".", it judges that it is an electronic mail address, and the finder application 54C activates the electronic mail program 56D, sets that electronic mail address, and carries out preparation of mail transmission.

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号
特開2000-207487
(P2000-207487A)

(43) 公開日 平成12年7月28日 (2000.7.28)

(51) Int.Cl. ⁷	識別記号	F I	テマコード [*] (参考)
G 0 6 K 9/03		G 0 6 K 9/03	J 5 B 0 6 4
G 0 6 F 3/00	6 5 4	G 0 6 F 3/00	6 5 4 D 5 E 5 0 1
G 0 6 K 9/00		G 0 6 K 9/00	S

審査請求 未請求 請求項の数6 O L (全 11 頁)

(21) 出願番号 特願平11-11953

(22) 出願日 平成11年1月20日 (1999.1.20)

(71) 出願人 000002185

ソニー株式会社

東京都品川区北品川6丁目7番35号

(72) 発明者 中嶋 信二

東京都品川区北品川6丁目7番35号 ソニー株式会社内

(72) 発明者 塩野 智樹

東京都品川区北品川6丁目7番35号 ソニー株式会社内

(74) 代理人 100082131

弁理士 稲本 義雄

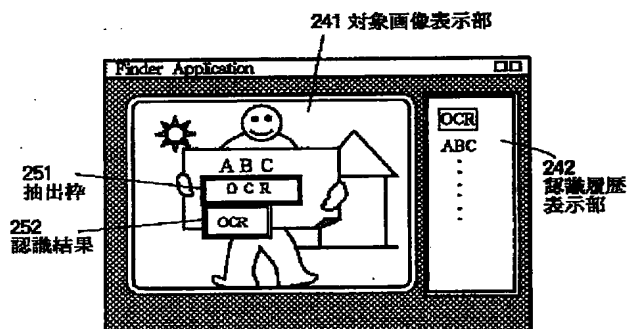
最終頁に続く

(54) 【発明の名称】 情報処理装置および方法、並びに提供媒体

(57) 【要約】

【課題】 パーソナルコンピュータに対して、簡単かつ確実に、所定の指令を入力させる。

【解決手段】 CCDビデオカメラにより所定のオブジェクトに表示されている文字を撮像させ、その撮像された画像を文字認識した結果得られた認識結果252を、対象画像表示部241に重ねて表示させる。認識結果252の近傍には、その認識に利用した画像データの範囲を示す抽出枠251が表示される。これにより、ユーザは、認識対象となる画像全体から、どの部分が抽出されて、認識結果が得られたのかを、直感的に把握することができる。



【特許請求の範囲】

【請求項 1】 被写体を撮像する撮像手段と、
前記撮像手段により撮像された画像を表示する対象画像
表示手段と、

前記撮像手段により撮像された画像のデータに基づい
て、文字認識処理を行う文字認識手段と、
前記文字認識手段による認識結果を表示する認識結果表
示手段と、
前記撮像手段により撮像された画像のうち、前記文字認
識手段による認識結果が得られた領域を表示する抽出領
域表示手段とを備えることを特徴とする情報処理装置。

【請求項 2】 前記認識結果表示手段は、前記対象画像
表示手段によって表示された画像上に、前記文字認識手
段による認識結果を表示させることを特徴とする請求項
1に記載の情報処理装置。

【請求項 3】 前記文字認識手段によって認識された認
識結果の履歴を順次表示する認識履歴表示手段をさらに
備えることを特徴とする請求項 1に記載の情報処理装
置。

【請求項 4】 前記第 3 の表示制御手段は、前記文字認
識手段による認識結果が得られた領域を、前記第 1 の表
示制御手段によって表示された画像上に、抽出枠として
表示させることを特徴とする請求項 1に記載の情報処理
装置。

【請求項 5】 被写体を撮像する撮像処理ステップと、
前記撮像処理ステップで撮像された画像を表示する対象
画像表示処理ステップと、
前記撮像処理ステップで撮像された画像のデータに基づ
いて、文字認識処理を行う文字認識処理ステップと、
前記文字認識処理ステップによる認識結果を表示する認
識結果表示処理ステップと、
前記撮像処理ステップで撮像された画像のうち、前記文
字認識処理ステップによる認識結果が得られた領域を表
示する抽出領域表示処理ステップとを備えることを特徴
とする情報処理方法。

【請求項 6】 被写体を撮像する撮像処理ステップと、
前記撮像処理ステップで撮像された画像を表示する対象
画像表示処理ステップと、
前記撮像処理ステップで撮像された画像のデータに基づ
いて、文字認識処理を行う文字認識処理ステップと、
前記文字認識処理ステップによる認識結果を表示する認
識結果表示処理ステップと、
前記撮像処理ステップで撮像された画像のうち、前記文
字認識処理ステップによる認識結果が得られた領域を表
示する抽出領域表示処理ステップとを含む処理を情報処
理装置に実行させるコンピュータが読み取り可能なプロ
グラムを提供することを特徴とする提供媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、情報処理装置およ

び方法、並びに提供媒体に関し、特に、撮像された画像
のデータから文字を認識し、所定の処理を行うことができ
るようにした情報処理装置および方法、並びに提供媒
体に関する。

【0002】

【従来の技術】物品の種類や状態などを示す英数文字を
コード化し、例えば、物品に貼付しておき、それを、後
で読み取って情報（物品の種類や状態など）を取得する
システム、いわゆる、バーコードシステムが、現在、多
くの産業分野に普及している。

【0003】図 1 は、バーコードの例を表している。こ
の例において、バーコードラベル 500 には、コード部
501 の下側の文字表示部 502 に表示されている英数
文字が、コード部 501 に配置されるバー（黒色の部
分）の幅と、各バーの間（白色の部分）の幅に基づいて
コード化されている。このように、情報を 1 つのパラメ
ータ（バーまたはバー間の幅）によりコード化するバー
コードは、1 次元バーコードと称される。

【0004】ところで、バーコードラベルから、バーコ
ードを読み取る 1 つの方法として、ビデオカメラにより
読み取る方法が上げられるが、例えば、コード化される
情報に伴って、バーの配列が多くなると、ビデオカメラ
によるバーコードの読み取りが困難になる。

【0005】そこで、コード化される情報が多くなった
場合においても、ビデオカメラによる読み取りを可能に
するために、1 次元バーコードのバーに代わる、図 2 に
示すようなセル（黒色の方形）が、2 次元的に配列され
た 2 次元バーコードが、提案されている。これによれ
ば、情報が 2 次元的にコード化されることより、より多
くの情報量をコード化することができ、1 次元バーコー
ドのバーに比べ、多くの情報がコード化されていても、
ビデオカメラによる読み取りが容易となる。

【0006】

【発明が解決しようとする課題】しかしながら、1 次元
バーコードおよび 2 次元バーコードのいずれも、そのよ
うなコードパターンを予め用意しておかなければなら
ず、一般ユーザが、例えばパーソナルコンピュータの各
種の処理を制御するために利用することができない課題
があった。

【0007】また、特開平 9-289624 号公報に
は、デジタルカメラに OCR (Optical Character Reade
r) を組み合わせ、カメラで取り込んだ画像を文字認識
することが提案されている。

【0008】しかしながら、文字認識は必ずしも常に正
しく行われるものではないので、認識された結果に基づ
いて、各種の処理を正確かつ確実に制御することが困難
である課題があった。

【0009】本発明はこのような状況に鑑みてなされた
ものであり、一般ユーザが、簡単かつ確実に、パーソナ
ルコンピュータなどの情報処理装置の処理を制御するこ

10

20

30

40

50

とができるようにするものである。

【0010】

【課題を解決するための手段】請求項1に記載の情報処理装置は、被写体を撮像する撮像手段と、撮像手段により撮像された画像を表示する対象画像表示手段と、撮像手段により撮像された画像のデータに基づいて、文字認識処理を行う文字認識手段と、文字認識手段による認識結果を表示する認識結果表示手段と、撮像手段により撮像された画像のうち、文字認識手段による認識結果が得られた領域を表示する抽出領域表示手段とを備えることを特徴とする。

【0011】請求項5に記載の情報処理方法は、被写体を撮像する撮像処理ステップと、撮像処理ステップで撮像された画像を表示する対象画像表示処理ステップと、撮像処理ステップで撮像された画像のデータに基づいて、文字認識処理を行う文字認識処理ステップと、文字認識処理ステップによる認識結果を表示する認識結果表示処理ステップと、撮像処理ステップで撮像された画像のうち、文字認識処理ステップによる認識結果が得られた領域を表示する抽出領域表示処理ステップとを備えることを特徴とする。

【0012】請求項6に記載の提供媒体は、被写体を撮像する撮像処理ステップと、撮像処理ステップで撮像された画像を表示する対象画像表示処理ステップと、撮像処理ステップで撮像された画像のデータに基づいて、文字認識処理を行う文字認識処理ステップと、文字認識処理ステップによる認識結果を表示する認識結果表示処理ステップと、撮像処理ステップで撮像された画像のうち、文字認識処理ステップによる認識結果が得られた領域を表示する抽出領域表示処理ステップとを含む処理を情報処理装置に実行させるコンピュータが読み取り可能なプログラムを提供することを特徴とする。

【0013】請求項1に記載の情報処理装置、請求項5に記載の情報処理方法、および請求項6に記載の提供媒体においては、撮像された画像から、文字認識が行われ、認識結果が表示されるとともに、認識結果が得られた領域が表示される。

【0014】

【発明の実施の形態】図3乃至図8は、本発明を適用した携帯型パーソナルコンピュータの構成例を表している。このパーソナルコンピュータ1は、ミニノート型のパーソナルコンピュータとされ、基本的に、本体2と、本体2に対して開閉自在とされている表示部3により構成されている。図3は、表示部3を本体2に対して開いた状態を示す外観斜視図、図4は、表示部3を本体2に対して開いた状態の平面図、図5は、表示部3を本体2に対して閉塞した状態を示す左側側面図、図6は、表示部3を本体2に対して180度開いた状態を示す右側側面図、図7は、表示部3を本体2に対して閉塞した状態の正面図、図8は、表示部3を本体2に対して開いた状

態の底面図である。

【0015】本体2には、各種の文字や記号などを入力するとき操作されるキーボード4、マウスカーソルを移動させるときなどに操作されるトラックポイント（商標）5が、その上面に設けられている。また、本体2の上面には、音を出力するスピーカ8と、表示部3に設けられているCCDビデオカメラ23で撮像するとき操作されるシャッターボタン10がさらに設けられている。

【0016】表示部3の上端部には、ツメ13が設けられており、図5に示すように、表示部3を本体2に対して閉塞した状態において、ツメ13に対向する位置における本体2には、ツメ13が嵌合する孔部6が設けられている。本体2の前面には、スライドレバー7が前面に平行に移動可能に設けられており、スライドレバー7は孔部6に嵌合したツメ13と係合してロックし、またロック解除することができるようになっている。ロックを解除することにより、表示部3を本体2に対して回動することができる。ツメ13の隣りには、マイクロホン24が取り付けられている。このマイクロホン24は、図8にも示すように、背面からの音も収音できるようになされている。

【0017】本体2の正面にはまた、プログラマブルパワーキー（PPK）9が設けられている。本体2の右側面には、図6に示すように、排気孔11が設けられており、本体2の前面下部には、図7に示すように、吸気孔14が設けられている。さらに、排気孔11の右側には、PCMCIA（Personal Computer Memory Card International Association）カード（PCカード）を挿入するためのスロット12が設けられている。

【0018】表示部3の正面には、画像を表示するLCD（Liquid Crystal Display）21が設けられており、その上端部には、撮像部22が、表示部3に対して回動自在に設けられている。すなわち、この撮像部22は、LCD21と同一の方向と、その逆の方向（背面の方向）との間の180度の範囲の任意の位置に回動することができるようになされている。撮像部22には、CCDビデオカメラ23が取り付けられている。

【0019】表示部3の下側の本体側には、電源ランプPL、電池ランプBL、メッセージランプML、その他のLEDよりなるランプが設けられている。なお、図5に示す符号40は、本体2の左側面に設けられた電源スイッチであり、図7に示す符号25は、CCDビデオカメラ23のフォーカスを調整する調整リングである。さらに、図8に示す符号26は、本体2内に増設メモリを取り付けるための開口部を被覆する蓋であり、符号41は、蓋26のロックツメを外すためのピンを挿入する小孔である。

【0020】図9は、パーソナルコンピュータ1の内部の構成例を表している。内部バス51には、CPU（Central Processing Unit）52、必要に応じて挿入されるPCカード53、RAM（Random Access Memory）54、およ

びグラフィックチップ81が接続されている。この内部バス51は、外部バス55に接続されており、外部バス55には、ハードディスクドライブ(HDD)56、I/O(入出力)コントローラ57、キーボードコントローラ58、トラックポイントコントローラ59、サウンドチップ60、LCDコントローラ83、モデム50等が接続されている。

【0021】CPU52は、各機能を統括するコントローラであり、PCカード53は、オプションの機能を付加するとき適宜装着される。

【0022】グラフィックチップ81には、CCDビデオカメラ23で取り込んだ画像データが、処理部82で処理された後、入力されるようになされている。グラフィックチップ81は、処理部82を介してCCDビデオカメラ23より入力されたビデオデータを、内蔵するVRAM81Aに記憶し、適宜、これを読み出して、LCDコントローラ83に出力する。LCDコントローラ83は、グラフィックチップ81より供給された画像データをLCD21に出力し、表示させる。バックライト84は、LCD21を後方から照明するようになされている。

【0023】RAM54の中には、起動が完了した時点において、例えば、OS(Operating System)(基本プログラム)54A、オートパイロットプログラム(アプリケーションプログラム)54B、そしてファインダアプリケーション(アプリケーションプログラム)54Cなど、所定のものが、HDD56から転送され、記憶される。

【0024】OS(基本プログラムソフトウェア)54Aは、Windows98(商標)に代表される、コンピュータの基本的な動作を制御するものである。

【0025】オートパイロットプログラム54Bは、予め設定された複数の処理(またはプログラム)等を、予め設定された順序で順次起動して、処理するプログラムである。

【0026】ファインダアプリケーション54Cは、CCDビデオカメラ23がとらえた画像をリアルタイムで受け取り、逐次その画像の2値化処理と文字領域抽出処理を行って、文字認識エンジン89に、その文字領域抽出処理された2値化画像を送ったり、CCDビデオカメラ23がとらえた画像、文字抽出領域、文字認識結果などをGUIを使って効果的に表示したり、認識結果に応じて、その他のアプリケーションプログラムへの連動処理などを行う。

【0027】一方、外部バス55側のハードディスクドライブ(HDD)56には、OS(基本プログラムソフトウェア)56A、オートパイロットプログラム56B、およびファインダアプリケーション56Cが記憶されている。

【0028】ハードディスクドライブ56には、この他、電子メールプログラム56D、インターネットブラ

ウザプログラム56E、ワードプロセッサプログラム56F、翻訳プログラム56Gなどのアプリケーションプログラムが記憶されている。

【0029】電子メールプログラム56Dは、電話回線のような通信回線等からネットワーク経由で通信文を受受するプログラムである。電子メールプログラム56Dは、特定機能としての着信メール取得機能を有している。この着信メール取得機能は、メールサーバ93に対してそのメールボックス93A内に自分(利用者)宛のメールが着信しているかどうかを確認して、自分宛のメールがあれば取得する処理を実行する。

【0030】インターネットブラウザプログラム56Eは、例えば、ネットスケープ社(商号)のネットスケープ(商標)のように、インターネット92を介して、他の情報処理装置と通信する機能を実行する。

【0031】ワードプロセッサプログラム56Fは、例えば、一太郎(商標)、ワード(商標)などのように、日本語あるいは英語などのワードプロセス処理を実行する。翻訳プログラム56Gは、日本語を英語に、あるいはその逆に、英語を日本語に翻訳する処理を実行する。

【0032】I/Oコントローラ57は、マイクロコントローラ61を有し、このマイクロコントローラ61にはI/Oインタフェース62が設けられている。このマイクロコントローラ61はI/Oインタフェース62、CPU63、RAM64、ROM69が相互に接続されて構成されている。このRAM64は、キー入力ステータスレジスタ65、LED(発光ダイオード)制御レジスタ66、設定時刻レジスタ67、レジスタ68を有している。設定時刻レジスタ67は、ユーザが予め設定した時刻(起動条件)になると起動シーケンス制御部76の動作を開始させる際に利用される。レジスタ68は、予め設定された操作キーの組み合わせ(起動条件)と、起動すべきアプリケーションプログラムの対応を記憶するもので、その記憶された操作キーの組み合わせがユーザにより入力されると、その記憶されたアプリケーションプログラム(例えば電子メールプログラム56D)が起動されることになる。

【0033】キー入力ステータスレジスタ65は、ワンタッチ操作作用のプログラマブルパワーキー(PPK)9が押されると、操作キーフラグが格納されるようになっている。LED制御レジスタ66は、レジスタ68に記憶されたアプリケーションプログラム(電子メールプログラム56D)の立上げ状態を表示するメッセージランプMLの点灯を制御するものである。設定時刻レジスタ67は、所定の時刻を任意に設定することができるものである。

【0034】なお、このマイクロコントローラ61にはバックアップ用のバッテリー74が接続されており、各レジスタ65、66、67の値は、本体2の電源がオフとされている状態においても保持されるようになっている。

10

20

30

40

50

る。

【0035】マイクロコントローラ61内のROM69の中には、ウェイクアッププログラム70、キー入力監視プログラム71、LED制御プログラム72が予め格納されている。このROM69は、例えばEEPROM (electrically erasable and programmable read only memory) で構成されている。このEEPROMはフラッシュメモリとも呼ばれている。さらにマイクロコントローラ61には、常時現在時刻をカウントするRTC (Real-Time Clock) 75が接続されている。

【0036】ROM69の中のウェイクアッププログラム70は、RTC75から供給される現在時刻データに基づいて、設定時刻レジスタ67に予め設定された時刻になったかどうかをチェックして、設定された時刻になると、所定の処理（またはプログラム）等の起動をするプログラムである。キー入力監視プログラム71は、PPK9が利用者により押されたかどうかを常時監視するプログラムである。LED制御プログラム72は、メッセージランプMLの点灯を制御するプログラムである。

【0037】ROM69には、さらにBIOS (Basic Input/Output System) 73が書き込まれている。このBIOSとは、基本入出力システムのことをいい、OSやアプリケーションソフトウェアと周辺機器（ディスプレイ、キーボード、ハードディスクドライブ等）の間でのデータの受け渡し（入出力）を制御するソフトウェアプログラムである。

【0038】外部バス55に接続されているキーボードコントローラ58は、キーボード4からの入力をコントロールする。トラックポイントコントローラ59はトラックポイント5の入力を制御する。

【0039】サウンドチップ60は、マイクロホン24からの入力を取り込み、あるいは内蔵スピーカ8に対して音声信号を供給する。

【0040】モデム50は、公衆電話回線90、インターネットサービスプロバイダ91を介して、インターネット等の通信ネットワーク92やメールサーバ93等に接続することができる。

【0041】電源スイッチ40は、電源をオンまたはオフするとき操作される。半押しスイッチ85は、シャッターボタン10が半押し状態にされたときオンされ、全押しスイッチ86は、シャッターボタン10が全押し状態にされたときオンされる。反転スイッチ87は、撮像部22が180度回転されたとき (CCDビデオカメラ23がLCD21の反対側を撮像する方向に回転されたとき)、オンされるようになされている。

【0042】次に、本発明のパーソナルコンピュータ1において、ファインダアプリケーション54Cが実行された場合の処理について説明する。このファインダアプリケーション54Cが起動されている場合、ユーザは、図10に示すように、所定の文字が印刷された厚紙など

で構成されるカード状のオブジェクト111を、CCDビデオカメラ23で撮像することで、パーソナルコンピュータ1の動作を制御することができる。ファインダアプリケーション54Cの処理について、図11および図12のフローチャートを参照して説明する。

【0043】最初にステップS1において、ファインダアプリケーション54Cは、グラフィックチップ81を制御し、GUI (Graphical User Interface) を作成し、表示させる。グラフィックチップ81は、ファインダアプリケーション54Cからの指令に対応して、GUIの画像データを作成し、VRAM81Aに描画する。VRAM81Aの画像データはそこから読み出され、LCDコントローラ83を介して、LCD21に出力され、表示される。これにより、例えば図13に示すようなファインダアプリケーションのGUIがLCD21に表示される。このGUIにおいては、CCDビデオカメラ23が撮像した画像を表示する対象画像表示部241と、後述するようにして文字認識エンジン89により文字認識がなされた結果の履歴を表示する認識履歴表示部242が設けられている。

【0044】次にステップS2において、ファインダアプリケーション54Cは、CCDビデオカメラ23（より具体的には、そのドライバ）から画像データを受信したか否かを判定し、受信していない場合には、受信するまで待機する。画像データを受信した場合、ステップS3に進み、ファインダアプリケーション54Cは、その画像をGUI上に表示させる。即ち、CCDビデオカメラ23により撮像されたオブジェクト111の画像は、処理部82により、所定の処理が施された後、グラフィックチップ81に供給される。グラフィックチップ81は、この画像データをVRAM81AのGUIの対象画像表示部241に対応する部分に描画する。このVRAM81Aに描画されている画像データが読み出され、LCDコントローラ83を介して、LCD21に表示されるので、図13に示すように、対象画像表示部241にCCDビデオカメラ23で撮像した画像が動画像として表示される。また、CCDビデオカメラ23で撮像した画像のデータは、文字認識のための領域（文字領域）を抽出するために、RAM54にも記憶される。

【0045】ステップS4において、ファインダアプリケーション54Cは、ユーザにより文字領域が指定されたか否かを判定する。文字領域が指定されていない場合には、ステップS5に進み、ファインダアプリケーション54Cは、RAM54に記憶されている1フレーム分の画像のうちの所定の範囲を文字領域として抽出する。そして、ステップS6において、ファインダアプリケーション54Cは、抽出された文字領域の画素データを文字認識エンジン89に転送し、文字認識させる。文字認識エンジン89は、各種の文字のパターンを予め記憶しており、そのパターンと転送されてきた文字領域の範囲の画素データとを比較し、文字が含まれているか否かを認

識する。そして、文字認識エンジン89は、認識した結果(文字)をファインダアプリケーション54Cに出力する。

【0046】このように、通常、ファインダアプリケーション54Cは、各フレームの画像に対する文字認識処理を、連続的に実行するが、ユーザが、トラックポイント5などを操作することで、文字領域を指定することも可能である。ユーザにより、文字領域が指定された場合には、ステップS5の処理はスキップされ、その指定された文字領域の範囲において、文字認識エンジン89により文字認識処理が実行される。

【0047】ステップS7において、ファインダアプリケーション54Cは、文字認識エンジン89の出力に基づいて、文字が認識されたか否かを判定し、認識されていない場合には、ステップS2に戻り、それ以降の処理を繰り返し実行する。

【0048】ステップS7において、文字が認識されたと判定された場合、ステップS8に進み、ファインダアプリケーション54Cは、文字が検出されたその文字領域の範囲を示す枠をGUIに表示させる。即ち、このとき、ファインダアプリケーション54Cは、グラフィックチップ81を制御し、GUIの対象画像表示部241内の文字が認識された文字領域を抽出枠251で表示させる。これにより、ユーザは、どこの領域の画像から文字が認識されたのかを知ることができる。さらに、ステップS9において、ファインダアプリケーション54Cは、文字認識エンジン82より供給された文字認識結果(文字)252を、抽出枠251の下に表示させる。なお、このとき、文字の認識結果として、複数の候補が存在する場合には、その候補も同時に表示される。ユーザは、トラックポイント5などを操作することで、文字認識結果252の中から正しい文字を選択したり、既に正しい文字が1つ表示されている場合には、その認識結果を確定させるための入力操作を行う。文字認識結果252の近傍に、それが認識された範囲を示す抽出枠251が表示されるので、ユーザは、その認識結果が正しいものであるか否かを、迅速かつ確実に判断することができる。

【0049】そこで、ステップS10において、ファインダアプリケーション54Cは、認識結果が正しいとの確定入力がなされたか否かを判定し、確定入力がなされない場合には、ステップS2に戻り、それ以降の処理を繰り返し、実行する。

【0050】正しい認識結果の確定入力がなされた場合には、ステップS11において、ファインダアプリケーション54Cは、RAM54またはハードディスクドライブ56内のエディットボックスと履歴ボックスに、確定された文字認識結果をコピーさせる。エディットボックスに登録された文字認識結果は、後に適宜これを読み出して利用することができる。

【0051】次に、ステップS12において、ファインダアプリケーション54Cは、履歴ボックス内の内容を読み出し、GUIの認識履歴表示部242に表示させる。その結果、認識履歴表示部242には、図13に示すように、文字認識結果の履歴が表示される(認識された文字が順番に表示される)。

【0052】ファインダアプリケーション54Cは、さらに、ステップS13において、認識履歴表示部242に表示されている文字認識結果の中から、所定のものが選択されたか否かを判定する。すなわち、ユーザは、認識履歴表示部242に表示されている文字の中から所定のものを、トラックポイント5などを操作することで選択することができる。所定の文字が選択されたとき、ステップS14において、ファインダアプリケーション54Cは、選択された文字に対応する処理を実行する。

【0053】例えば、ファインダアプリケーション54Cは、選択された文字列が「http://」で始まる文字列であれば、それをURL(Uniform Resource Locator)と判断し、インターネットブラウザプログラム56Eを起動し、そのURLにアクセスさせる。選択された文字列が、「@」と「.」とで区切られた文字列である場合には、電子メールアドレスであると判断し、ファインダアプリケーション54Cは、電子メールプログラム56Dを起動し、その電子メールアドレスを設定し、メール送信の準備を行う。

【0054】その文字列がハードディスクドライブ56に予め登録されている各種のアプリケーションプログラムの名称である場合には、ファインダアプリケーション54Cは、その名称のアプリケーションプログラムを起動させる。アプリケーションプログラムの名称の後に、さらに文字列がある場合には、その文字列をオプションあるいは引き数として、アプリケーションプログラムが起動される。例えば、その文字列にワードプロセッサプログラム56Fの名称が含まれている場合には、ファインダアプリケーション54Cは、ワードプロセッサプログラム56Fを起動する。そして、アプリケーションプログラムの名称の後に、ファイル名を示す文字列が認識された場合は、起動したワードプロセッサプログラム56Fに対して、そのファイルを開く処理を指示する。

【0055】選択された文字列が一般的な単語である場合には、ファインダアプリケーション54Cは、翻訳プログラム56Gを起動し、その文字列を翻訳させる。

【0056】従って、ユーザは、所定の文字が表示されているオブジェクト111をCCDビデオカメラ23により撮像させることで、パーソナルコンピュータ1に所定の動作を実行させることができる。

【0057】ステップS13において、文字認識結果の選択が行われていないと判定された場合、ステップS14の処理はスキップされる。その後、ステップS2に戻り、それ以降の処理が繰り返し実行される。

【0058】以上においては、本発明を携帯型パーソナルコンピュータに応用した場合を例として説明したが、本発明は、デスクトップ型のパーソナルコンピュータ、その他の情報処理装置にも応用することが可能である。

【0059】なお、上記したような処理を行うコンピュータプログラムをユーザに提供する提供媒体としては、磁気ディスク、CD-ROM、固体メモリなどの記録媒体の他、ネットワーク、衛星などの通信媒体を利用することができる。

【0060】

【発明の効果】以上の如く、請求項1に記載の情報処理装置、請求項5に記載の情報処理方法、および請求項6に記載の提供媒体によれば、認識対象となる画像と、文字認識された認識結果と、この認識結果が得られた抽出領域を同時に表示するようにしたので、ユーザは、認識対象となる画像全体から、どの部分が抽出されて認識結果が得られたのかを直感的に把握することができ、さらに、情報処理装置に対して所定の指令を入力し、所定の処理を実行させることが可能となる。

【図面の簡単な説明】

【図1】1次元バーコードの例を示す図である。

【図2】2次元バーコードの例を示す図である。

【図3】本発明を適用した携帯型パーソナルコンピュータの構成例を示す斜視図である。

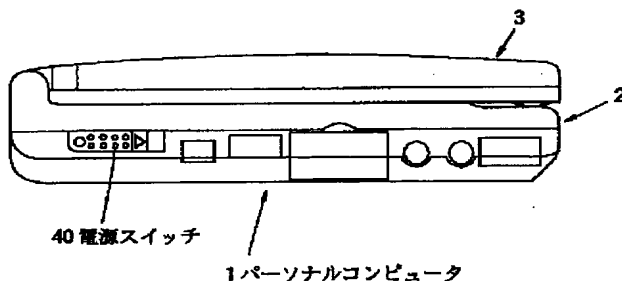
【図4】図3の携帯型パーソナルコンピュータの表示部を開いた状態の平面図である。

*

【図1】



【図5】



* 【図5】図3の携帯型パーソナルコンピュータの表示部を閉じた状態の左側面図である。

【図6】図3の携帯型パーソナルコンピュータの表示部を開いた状態の右側面図である。

【図7】図3の携帯型パーソナルコンピュータの表示部を閉じた状態の正面図である。

【図8】図3の携帯型パーソナルコンピュータの表示部を開いた状態の底面図である。

10 【図9】図3の携帯型パーソナルコンピュータの内部の構成例を示すブロック図である。

【図10】図3の携帯型パーソナルコンピュータの使用状態を示す斜視図である。

【図11】図3の携帯型パーソナルコンピュータの動作を説明するフローチャートである。

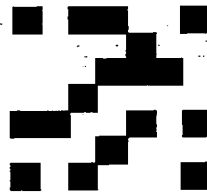
【図12】図3の携帯型パーソナルコンピュータの動作を説明するフローチャートである。

【図13】GUIの表示例を示す図である。

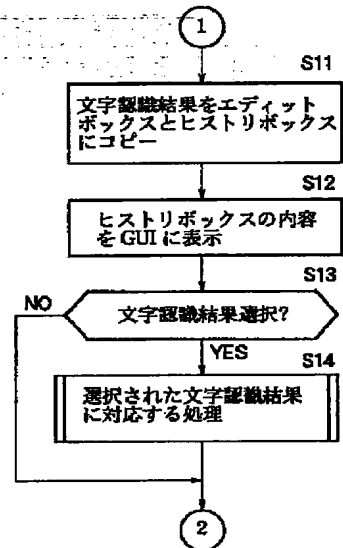
【符号の説明】

- 1 携帯型パーソナルコンピュータ, 2 本体, 3 表示部, 21 LCD, 23 CCDビデオカメラ, 52 CPU, 54 RAM, 54 C ファインダアプリケーション, 56 D 電子メールプログラム, 56 E インターネットブラウザプログラム, 56 F ワードプロセッサプログラム, 56 G 翻訳プログラム, 89 文字認識エンジン

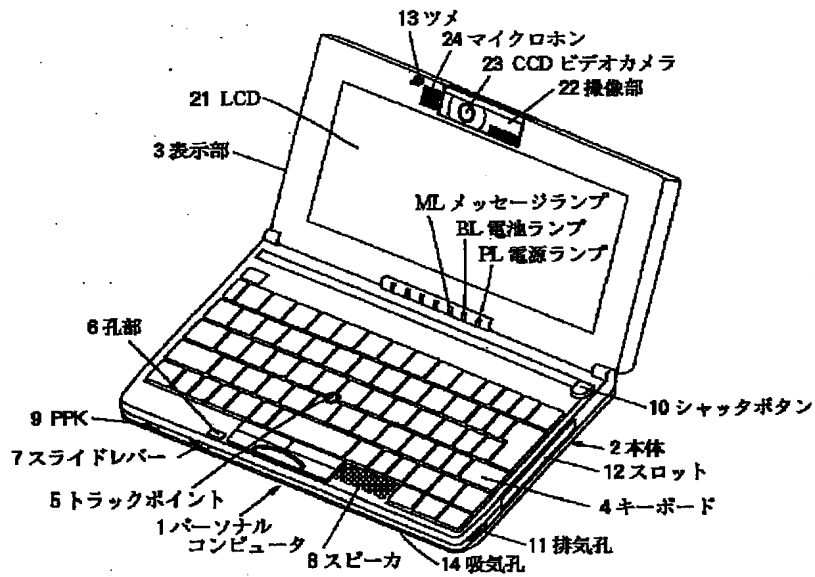
【図2】



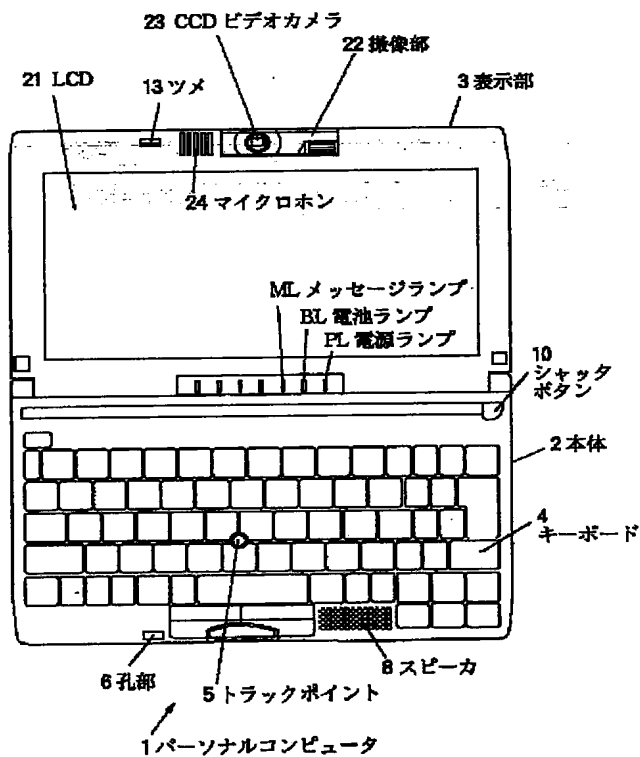
【図12】



【図3】



【図4】



【図8】

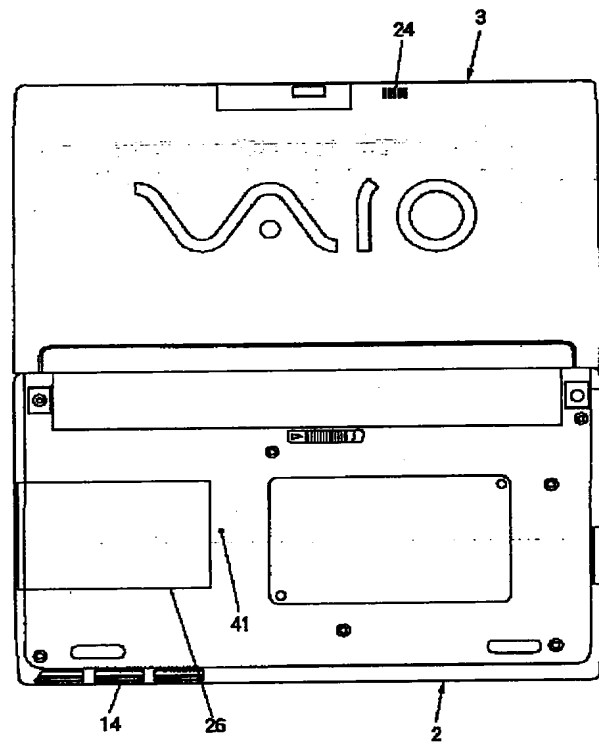


Fig. 10

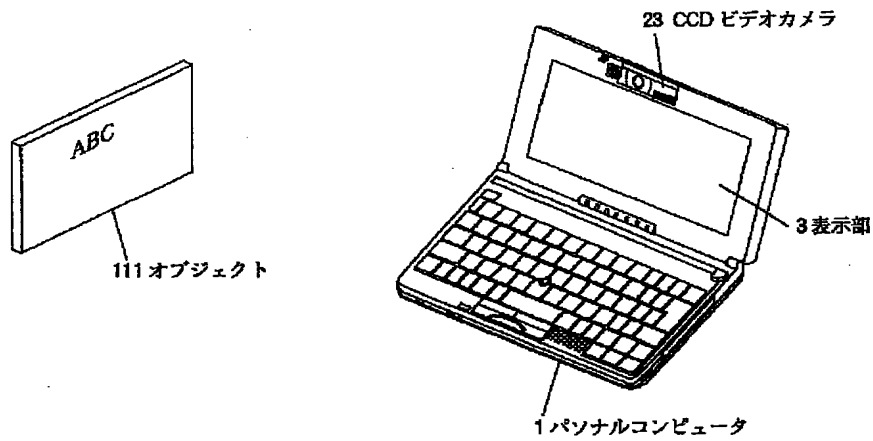
11 排気孔
12 スロット

Figure 1 is a perspective view of the front of the camera body. The camera is shown from a front-three-quarter angle. The lens barrel is on the left, and the viewfinder and various controls are on the right. The camera body is labeled with the number 1. The lens barrel is labeled with the number 2. The viewfinder is labeled with the number 3. The controls on the right are labeled with the number 4. The camera body is labeled with the number 1. The lens barrel is labeled with the number 2. The viewfinder is labeled with the number 3. The controls on the right are labeled with the number 4.

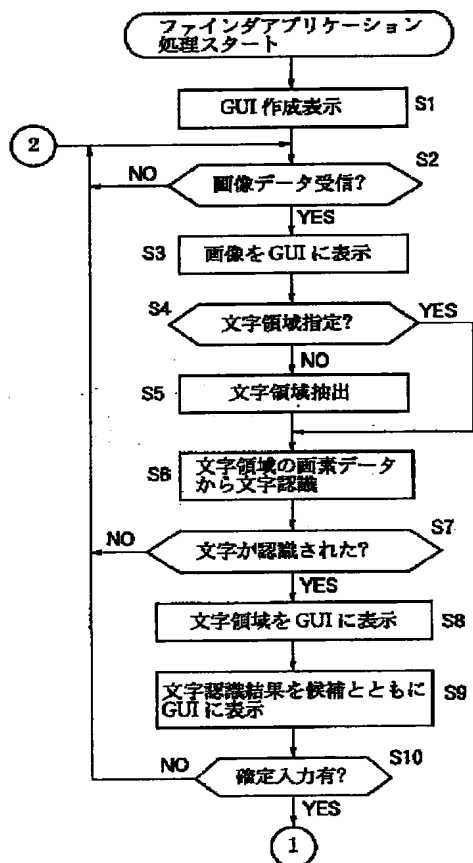
Figure 1 is a block diagram of a portable electronic device (1). The device includes a main body (1) and a power-on sequence control unit (75). Key components and their connections are as follows:

- Input/Output (I/O) Interface (62):** Connects to a network (92) via a modem interface (90) and a network interface (91). It also connects to a modem (90) and a network (92) via a modem interface (90).
- Internal Components:**
 - CPU (83):** The central processing unit.
 - RAM (64):** Random Access Memory.
 - ROM (65):** Read-Only Memory.
 - Registers (66, 67):** Various control registers.
 - Display (LCD 21):** Liquid Crystal Display.
 - Keyboard (4):** Input device.
 - Microphone (24):** Audio input.
 - Speaker (8):** Audio output.
 - Battery (74):** Power source.
 - Power Switch (40):** Main power control.
 - Reset Switch (88):** System reset control.
- External Connections:**
 - Modem (90):** Connected to the network (92) and the device's I/O interface.
 - Network (92):** External network connection.
 - Modem Interface (91):** Connects the device to the modem.
- Power-On Sequence Control Unit (75):** Manages the device's startup sequence, including the power switch (40) and reset switch (88).

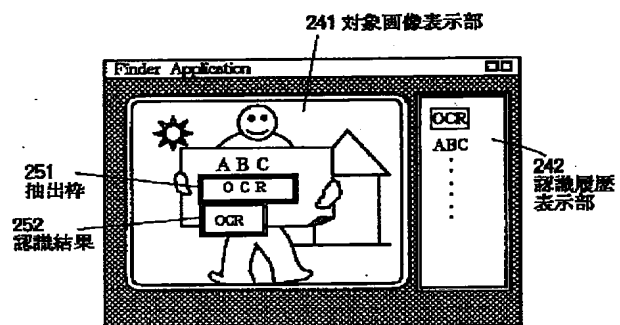
【図10】



【図11】



【図13】



フロントページの続き

(72)発明者 末吉 隆彦
東京都品川区北品川 6 丁目 7 番 35 号 ソニ
ー株式会社内

F ターム (参考) 5B064 BA01 EA11 EA12 FA04 FA06
FA13
5E501 AA03 AA14 AC16 BA05 CA04
CB08 CB14 DA17 EA13 EA34
EB01 EB05 EB06 EB19 FA15
FB46